

Appln. No. 09/775,285  
Response dated May 10, 2004  
Reply to Office Action of Mar. 9, 2004  
Docket No. 6169-149

IBM Docket No. BOC9-2000-0004

### REMARKS/ARGUMENTS

These remarks are made in response to the final Office Action of March 9, 2004 (Office Action). As this response is timely filed within the three-month shortened statutory period, no fee is due.

In paragraph 2 of the Office Action, claims 1-5, 6, and 9-18 have been rejected under 35 U.S.C. § 102(b) as being anticipated by U.S. Patent No. 6,088,671 to Gould *et al.* (Gould). In paragraph 15, claims 8 and 20 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Gould in view of U.S. Patent No. 6,539,080 to Bruce *et al.* (Bruce). In paragraph 17, claims 7 and 19 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Gould.

The Applicants respectfully traverse the above rejections. Additionally, the Applicants respectfully note that Gould was patented on July 11, 2000, and the present application was filed on February 1, 2001. As the present invention was filed less than one year after the patenting of Gould, it appears as though the 35 U.S.C. § 102(b) rejection with respect to claims 1-5, 6, and 9-18 is improper. The Applicants are proceeding, however, under the assumption that a different subsection of 35 U.S.C. § 102 was intended.

Prior to addressing the rejections on the art, a brief review of the Applicants' invention is in order. The Applicants have developed a method, system, and apparatus for presenting database query results to a user in an auditory fashion. As disclosed, each choice extracted from a database query is audibly presented immediately upon its extraction rather than following the conventional process of extracting all choices and then playing them in batch. The user can respond to each choice when that choice is presented, thereby interrupting the database query and the subsequent presentation of additional choices through an audio user interface (AUI).

In reviewing the Examiner's response, it appears, that the differences between database querying and presentation of results through an AUI, as described in the Applicants' specification, and speech recognition, as taught by Gould, have not been appreciated or have been overlooked. This view seems to pervade the Office Action. As such, the Applicants would like to clarify the nature of database query operations and speech recognition in general.

Speech recognition is performed to convert user spoken utterances to text. While recognized text may specify a query, in the case of Gould, the text is either dictation or a

Appln. No 09/775,285  
Response dated May 10, 2004  
Reply to Office Action of Mar 9, 2004  
Docket No. 6169-149

IBM Docket No BOC9-2000-0004

command. In contrast, a database query operation, as used in the instant application, whether initiated as a user spoken utterance or through some other means such as touch tone input, is processed by an information processing system to retrieve user requested information from a database. The query operation is distinct from, and wholly unrelated to, speech recognition.

As an example, consider a user requesting the telephone number of an individual through an automated directory assistance system. The user provides a speech input requesting a telephone number for an individual. Notably, the user can utilize other means of requesting the information such as touchtone input. In any case, the user speech is speech recognized to obtain a textual representation of the user input. Once processed, the user's request can be addressed. The system, i.e. an information processing, then can query a database to retrieve the information requested by the user, in this case the telephone number of an individual. Appreciably, the act of locating and finding the telephone number through a database query is distinct from the process of speech recognizing the user input. This distinction between speech recognition and database query operations is exemplified by the discussion on page 7, lines 18-29 of the Applicants' specification.

Turning to the rejections on the art, claims 1-5, 6, and 9-18 have been rejected under 35 U.S.C. § 102(b) as being anticipated by Gould. Gould is directed to a speech recognition system that can distinguish between dictation and commands. More particularly, the cited portions of Gould describe a speech recognition system for determining whether an utterance in a user speech input stream should be interpreted as dictated text or a command. Gould fails to teach or suggest database query operations or presenting query results via an AUI as explicitly claimed by the Applicants.

Regarding claims 1, 13, and point 1 in the Response to Argument portion of the Office Action, column 4, lines 49-67 have been cited for the proposition that Gould teaches presenting database query results. The cited passage, however, describes how Gould recognizes received speech input as text. In reference to the above discussion, database querying is wholly unrelated to speech recognition. Similarly, the discussion at column 6, lines 18-34 describes the process of how Gould recognizes user spoken utterances as speech or commands. It seems that Gould's discussion of comparing recognition results to command templates has been construed to be the

Appln No. 09/775,285  
Response dated May 10, 2004  
Reply to Office Action of Mar. 9, 2004  
Docket No. 6169-149

IBM Docket No. BOC9-2000-0004

same as performing a database query operation. Again, speech recognition is neither the same as, nor analogous to, database querying. Moreover, nowhere within the cited passages are the words database or query even mentioned.

Regarding point 3, it is asserted that Gould teaches an AUI through which a user inputs a query. In support, column 1, lines 9-32 and column 4, lines 12-15 have been cited. Further, it is asserted that the Applicants' claim does not specify that results of the query must be presented in an auditory manner.

Regarding the cited passages of Gould, again Applicants reiterate that speech recognition and database query operations are not the same. Gould discloses only that user dictation and commands can be received. Dictation is merely text to be recognized. A command instructs the speech recognition system to take an action other than merely recognizing and providing text to an application program. Neither is analogous to performing a database query operation where information requested by the user is retrieved from a database.

In any case, contrary to the Examiner's assertion, the Applicants' claims do in fact specify that query results are presented in an auditory manner. The Applicants' claims state "presenting each said query result item through the AUI as each said query result item is found in said at least one database, said presenting step occurring concurrently with said database query operation." The claim states that query results are presented through an audio user interface. At page 1, lines 6-14, an AUI has been described as follows:

Designing effective methods of presenting information to users can be challenging for application developers. In a non-visual environment, such as telephony, this problem can be complicated even further. Specifically, in an audio user interface ("AUI"), visual feedback is not available to cue users or to confirm the accurate contextual understanding of user speech input. Additionally, listening to and comprehending spoken information typically consumes more time and can be less comfortable to users when compared to the visual comprehension of the same information. The problems associated with an AUI can become exacerbated when text-to-speech technology is used to present information.

Consequently, the Applicants' claims explicitly require the presentation of query results in an auditory manner, i.e. through an AUI.

Appl. No. 09/775,285  
Response dated May 10, 2004  
Reply to Office Action of Mar. 9, 2004  
Docket No. 6169-149

IBM Docket No. BOC9-2000-0004

Moreover, Gould utterly fails to teach or suggest that query results can be presented through an AUI. Gould teaches only that user dictation can be visually displayed upon a display or that user speech commands can be acted upon by the speech recognition system. Gould neither queries for information requested by the user nor presents information to the user in an audible format.

Regarding claims 2, 14, and point 2, it is asserted that Gould teaches terminating a database query operation responsive to a speech response. In support, column 6, lines 24-34 have been cited. The Office Action contends that the cited passage illustrates that "when a match is made with a user's speech, the CPU is finished, and termination occurs." Here, Gould states that if user speech that is hypothesized to be dictation is ultimately determined to be dictation rather than a speech command, the CPU is finished.

The processing described by Gould is not terminated prior to completion of a query operation responsive to the detection of a speech response as the Applicants' claims suggest. Rather, the CPU of Gould is finished only because the user has stopped speaking and no further data remains to be processed. The processing is not terminated, but stops of its own accord. Notably, were another user input to be received, as claimed by the Applicants, Gould would not stop processing. In that case, Gould would continue processing, which action is contrary to claims 2 and 14.

In sum, the language of claims 2 and 14 requires the detection of a speech response selecting a query result and then terminating a database query operation responsive to that detection. Gould fails to teach or suggest (1) a database query operation, (2) the presentation of query results through an AUI, (3) the detection of a user response selecting a query result, or (4) the termination of a database query result.

Regarding claim 9, the Applicants fail to understand how Gould can teach a database manager for managing database query operations and a dialog manager for managing the presentation of database query results through an AUI as Gould fails to teach or suggest database querying, the presentation of query results through an AUI, or even an AUI. In any case, regarding point 6, column 3, lines 36-49 have been cited for the proposition that Gould teaches a

Appl. No. 09/775,285  
Response dated May 10, 2004  
Reply to Office Action of Mar 9, 2004  
Docket No. 6169-149

IBM Docket No. BOC9-2000-0004

dialog manager for managing the presentation of database query results through an AUI concurrently with a database query operation.

The cited passage, however, teaches how speech input (speech packers) are monitored to determine whether there is any input audio that needs to be processed for the purpose of speech recognition. The paragraph has nothing to do with database queries or the presentation of query results through an AUI. The comment that "[s]ince the system must recognize queries/input as they are being processed, there is a manager (monitor software) that manages what has been entered but not yet processed" is not on point with the Applicants' claims. This reasoning completely overlooks the functionality that is explicitly stated in the claims. A generalized software monitor function, one that relates to input processing in a speech recognition system, is not the same as or analogous to the database manager or dialog manager of the present invention.

Regarding claim 10, the Examiner asserts that Gould teaches a text-to-speech processor. In support, column 4, lines 12-15 have been cited. At column 4, lines 12-15, Gould states that:

As an alternative to dictating directly to an application, the user dictates text to a speech recognizer window, and after dictating a document, the user transfers the document (manually or automatically) to the application.

From the above passage, the Applicants fail to see how Gould "very explicitly teaches a text-to-speech processor" as stated in the Office Action. The above passage has nothing to do with a text-to-speech processor. This passage states only that a user can dictate to an application (speech-to-text) or the speech recognition system itself. Text-to-speech is a process whereby text is converted into speech, whether synthetic or recorded. Text-to-speech can be used in the context of an AUI, for example in telephony. As such, Gould fails to teach or suggest a text-to-speech processor for converting database query results into audible speech.

Regarding claim 11, a "barge-in" facility describes a sort of duplex mechanism where a user input can be received and processed despite the fact that the user spoke or provided input while the interactive system was playing an audio prompt to the user. The user effectively interrupts or "barges-in" on the audio output of the system. Still, the system understands that the user has responded and begins processing the user response. Systems without barge-in simply ignore the fact that a user response was received. That is, such systems do not monitor for a user

Appln. No. 09/775,285  
Response dated May 10, 2004  
Reply to Office Action of Mar. 9, 2004  
Docket No. 6169-149

IBM Docket No. BOC9-2000-0004

response until the system has finished playing the audible prompt, thereby requiring the user to repeat his or her input after the audio prompt has finished.

The Applicants note that as Gould does not provide results in an audible fashion, Gould has no need and, in fact, is utterly silent with respect to a barge-in facility. The passage cited in support of such as function (column 6, lines 24-34 in point 8) has been discussed already and bears no relation to AUI's or a barge-in facility. As noted, this passage discusses how Gould hypothesizes whether initially processed user speech is dictation or a speech command.

In light of the above, withdrawal of the 35 U.S.C. § 102(b) rejection with respect to claims 1-5, 6, and 9-18 is respectfully requested.

In paragraph 18, claims 8 and 20 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Gould in view of Bruce. The Examiner concedes that Gould does not explicitly teach that the AUI is a telephony interface, but contends that Bruce teaches such a limitation. In consequence, the Examiner asserts that it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the teachings of Gould with Bruce's step of incorporating a telephony interface into an AUI speech and voice recognition system in order to create a system where information can be transmitted from two geographically removed points.

Bruce, however, fails to cure the deficiencies of Gould. While Bruce teaches that a telephone can be used as a user interface to a networked computer system, Bruce, like Gould, fails to teach or suggest that database query results can be presented through an AUI as the results are determined concurrently with the execution of the database operation. Regarding point 10, at column 2, lines 54-63, Bruce describes how results may be transmitted to a driver. Bruce does not, however, teach or suggest that the step of presenting query items occurs concurrently with the database query operation. As such, neither Gould, Bruce, nor any combination thereof teach or suggest the Applicants' invention as claimed. Accordingly, withdrawal of the 35 U.S.C. § 103(a) rejection with respect to claims 8 and 20 is respectfully requested.

In paragraph 20, claims 7 and 19 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Gould. The Examiner concedes that Gould does not teach that the data

Appin. No. 09/775,285  
Response dated May 10, 2004  
Reply to Office Action of Mar. 9, 2004  
Docket No. 6169-149

IBM Docket No. BOC9-2000-0004


structure can be a stack or database. The Examiner has taken Official Notice, however, that databases and stacks are well known in the art. Therefore, the Examiner contends that it would have been obvious to one of ordinary skill in the art at the time of the invention to include the possibility of using a stack or database data structure in order to provide a variety of storage possibilities based on need.

The Applicants respectfully note that the Official Notice taken with respect to data structures for claims 7 and 19 fails to cure the deficiencies of Gould already discussed at length. Accordingly, withdrawal of the 35 U.S.C. § 103(a) rejection with respect to claims 7 and 19 is respectfully requested.

The Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. The Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Date: 5/10/04

  
\_\_\_\_\_  
Gregory A. Nelson, Registration No. 30,577  
Kevin T. Cuenot, Registration No. 46,283  
Brian K. Buchheit, Registration No. 52,667  
AKERMAN SENTERFITT  
Post Office Box 3188  
West Palm Beach, FL 33402-3188  
Telephone: (561) 653-5000